

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 06-051800

(43)Date of publication of application : 25.02.1994

(51)Int.Cl.

G10L 9/18

G10L 9/00

(21)Application number : 04-223227

(71)Applicant : SONY CORP

(22)Date of filing : 30.07.1992

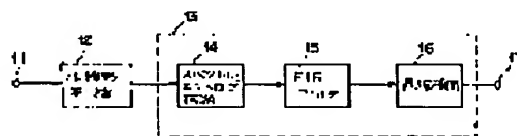
(72)Inventor : NISHIGUCHI MASAYUKI
MATSUMOTO ATSUSHI

(54) DATA QUANTITY CONVERTING METHOD

(57)Abstract:

PURPOSE: To convert the number of data while reducing an arithmetic quantity by using a coefficient set corresponding to each position of a certain number of data at a sample point of input data.

CONSTITUTION: A nonlinear compression part 12 performs the nonlinear compression of a variable number M of input data in blocks and a spectrum envelope extension part 14 repeatedly extends values at both ends of a spectrum envelope to the front and rear. The spectrum envelope whose ends are both extended to the front and rear is supplied to an FIR filter 15. This FIR filter 15 uses the coefficient set corresponding to a position nearby each position of a certain number of data among plural coefficient sets corresponding to mutually different phases at the sample point of the input data to find intermediate output data. The intermediate output data are supplied to a linear interpolation part 16 and linearly interpolated to obtain a certain number of data required for a final output, so that the data are outputted from an output terminal 17.



LEGAL STATUS

[Date of request for examination]

30.07.1999

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平6-51800

(43)公開日 平成 6 年(1994) 2月25日

(51)Int.Cl. ⁵	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L	9/18	E 8946-5H		
	9/00	M 8946-5H		

審査請求 未請求 請求項の数 4 (全 22 頁)

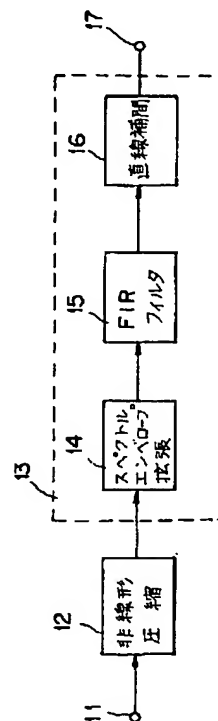
(21)出願番号	特願平4-223227	(71)出願人	000002185 ソニー株式会社 東京都品川区北品川 6 丁目 7 番35号
(22)出願日	平成 4 年(1992) 7月30日	(72)発明者	西口 正之 東京都品川区北品川 6 丁目 7 番35号 ソニー株式会社内
		(72)発明者	松本 淳 東京都品川区北品川 6 丁目 7 番35号 ソニー株式会社内
		(74)代理人	弁理士 小池 晃 (外 2 名)

(54)【発明の名称】 データ数変換方法

(57)【要約】

【構成】 入力されたブロック毎に可変個数のデータを、非線形圧縮部 1 2 で非線形圧縮し、スペクトルエンベロープ拡張部 1 4 でそのスペクトルエンベロープの両端を拡張し、F I R フィルタ 1 5 で F I R フィルタリング (演算) し、直線補間 1 6 で直線補間し、一定個数のサンプルデータに変換する。

【効果】 演算量を大幅に減らしながらも、データ個数変換ができる。



【特許請求の範囲】

【請求項1】 ブロック内の波形データ又は波形を表すパラメータデータの個数が可変とされたデータをブロック毎に一定の個数の基準データと比較するために上記可変個数の入力データを上記一定個数に変換する工程を含むデータ数変換方法において、

上記可変個数のデータが入力される帯域制限型オーバーサンプリングのためのFIRフィルタで、上記入力データのサンプル点に対してそれぞれ異なる複数の位相と対応した複数の係数セットの内の上記一定個数のデータの各位置に対応する係数セットを用いることにより、出力として必要な上記一定個数のデータのみを求める工程を有することを特徴とするデータ数変換方法。

【請求項2】 ブロック内の波形データ又は波形を表すパラメータデータの個数が可変とされたデータをブロック毎に一定の個数の基準データと比較するために上記可変個数の入力データを上記一定個数に変換する工程を含むデータ数変換方法において、

上記可変個数のデータが入力される帯域制限型オーバーサンプリングのためのFIRフィルタで、上記入力データのサンプル点に対してそれぞれ異なる複数の位相と対応した複数の係数セットの内の上記一定個数のデータの各位置の近傍の位置に対応する係数セットを用いることにより、中間的な出力データを求める工程と、

上記中間的な出力データを補間して最終出力に必要とされる一定個数のデータを求める工程とを有することを特徴とするデータ数変換方法。

【請求項3】 ブロック内の波形データ又は波形を表すパラメータデータが一定個数とされたデータをブロック内のデータ内容に応じた個数のデータに変換するデータ数変換方法において、

上記一定個数のデータが入力される帯域制限型オーバーサンプリングのためのFIRフィルタで、上記入力データのサンプル点に対してそれぞれ異なる複数の位相と対応した複数の係数セットの内の上記ブロック内のデータ内容に応じた個数のデータの各位置に対応する係数セットを用いることにより、出力として必要な上記ブロック内のデータ内容に応じた個数のデータのみを求める工程を有することを特徴とするデータ数変換方法。

【請求項4】 ブロック内の波形データ又は波形を表すパラメータデータが一定個数とされたデータをブロック内のデータ内容に応じた個数のデータに変換するデータ数変換方法において、

上記一定個数のデータが入力される帯域制限型オーバーサンプリングのためのFIRフィルタで、上記入力データのサンプル点に対してそれぞれ異なる複数の位相と対応した複数の係数セットの内の上記ブロック内のデータ内容に応じた個数のデータの各位置の近傍の位置に対応する係数セットを用いることにより、中間的な出力データを求める工程と、

上記中間的な出力データを補間して最終出力に必要とされる上記ブロック内のデータ内容に応じた個数のデータを求める工程とを有することを特徴とするデータ数変換方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、データ数変換方法に関し、特に、音声合成分析装置（ボコーダ）等において算出されたスペクトルの振幅データのような可変個数のデータを一定個数のデータに変換するようなデータ数変換方法に関する。

【0002】

【従来の技術】 オーディオ信号（音声信号や音響信号を含む）の時間領域や周波数領域における統計的性質と人間の聴感上の特性を利用して信号圧縮を行うような符号化方法が種々知られている。この符号化方法としては、大別して時間領域での符号化、周波数領域での符号化、分析合成符号化等が挙げられる。

【0003】 音声信号等の高能率符号化の例として、MBE（Multiband Excitation: マルチバンド励起）符号化、SBE（Singleband Excitation: シングルバンド励起）符号化、ハーモニック（Harmonic）符号化、SBC（Sub-band Coding: 帯域分割符号化）、LPC（Linear Predictive Coding: 線形予測符号化）、あるいはDCT（離散コサイン変換）、MDCT（モデファイドDCT）、FFT（高速フーリエ変換）等において、スペクトル振幅やそのパラメータ（LSPパラメータ、 α パラメータ、 k パラメータ等）のような各種情報データを量子化する場合に、従来においてはスカラー量子化を行うことが多い。

【0004】

【発明が解決しようとする課題】 ところで、ビットレートを例えば3～4 kbps程度にまで低減し、量子化効率を更に向上させようとする、スカラー量子化では量子化雑音（歪み）が大きくなってしまい、実用化が困難であった。そこで、これらの符号化の際に得られる時間軸データや周波数軸データやフィルタ係数データ等を個々に量子化せず、複数のデータを組（ベクトル）にまとめて一つの符号で表現して量子化するベクトル量子化が注目されている。

【0005】 しかしながら、上記MBE、SBE、LPC等のスペクトル振幅データ等は、ピッチに依存して個数が増えるため、そのままベクトル量子化しようすると可変次元のベクトル量子化が必要となり、構成が複雑化するのみならず、良好な特性を得ることが困難である。

【0006】 また、量子化の前にデータのブロック（フレーム）間差分をとるような場合にも、前後のブロック（フレーム）内のデータの個数が一致していないと、差分をとることができない。このように、可変個数のデー

タを一定個数に変換することがデータ処理の過程で必要とされることがあるが、特性の良好なデータ数変換が望まれる。

【0007】そこで、本出願人は、特願平4-92263号特許出願の明細書及び図面において、可変個数のデータを一定個数に変換することができ、端点でリンク等の発生しない特性の良好なデータ数変換が行えるようなデータ数変換方法を提案した。この方法は、ブロック毎に可変個数のデータを非線形圧縮部で非線形圧縮し、ダミーデータ付加部でブロック内の最後のデータ値から最初のデータ値までの補間をするようなダミーデータを付加してデータ個数を拡大した後、高速フーリエ変換（FFT）処理部、逆高速フーリエ変換（IFFT）処理部等を有した帯域制限型のオーバーサンプリング部でオーバーサンプルし、直線補間部で直線補間し、間引き処理部で間引くことにより一定個数のサンプルデータに変換するものである。

【0008】この出願によるデータ数変換方法では、FFTをする際に、1ブロックを例えば256サンプルに延長して計算している。次に、例えば8倍のオーバーサンプリングを実現するために、FFT変換により得られた256サンプルのスペクトルデータに対し、各サンプルの中間に7（ $=8-1$ ）個の0を詰めるような中間0詰め処理を行って2048サンプルとし、この2048サンプルに対してIFFTの計算を行っている。

【0009】ところで、通常のFFT、IFFTでは、1ブロックのサンプル数をNとすると、 $(N/2 \times \log_2 N)$ の複素乗算と、 $(N \log_2 N)$ の複素加算が行われている。ここで、 $(N/2 \log_2 N)$ の複素乗算は、 $(N/2 \times \log_2 N \times 4)$ の実数乗算となり、 $(N \log_2 N)$ の複素加算は、 $(N \log_2 N \times 2)$ の実数加算となる。したがって、Nを256としたときのFFTの演算量は、4096回（ $=256/2 \times 8 \times 4$ ）となり、N=2048としたときのIFFTの演算量は、45056回（ $=2048/2 \times 11 \times 4$ ）となり、その合計は49152回となる。

【0010】また、全実数入力に対して、N/2点のFFTで、N点FFTが実現できる、いわゆる高速化の手法を用いたとしても、 $N/4 (\log_2 N - 1) \times 4 + N \times 4$ の実数乗算と、 $N/2 (\log_2 N - 1) \times 2 + N \times 2$ の実数加算が必要となる。すなわち、N=256としたときのFFTでは、乗算が2816回、加算が2304回行われる。また、N=2048としたときのIFFTでは、乗算が28672回、加算が24576回行われる。したがって、乗算だけでも31488回の演算が必要となる。

【0011】なお、以上はエンコードの際において、ブロック（フレーム）内で可変個数（8～63個）のサンプルデータを一定個数（44個）のサンプルデータに変換するデータ数（サンプルレート）変換を想定している

が、デコードの場合も同様な方法でブロック（フレーム）内の一定個数（44個）のサンプルデータを可変個数（8～63個）のサンプルデータに変換しているものである。

【0012】ところで、実際に求めたい点の数は、エンコードの際には2048点でIFFTした内の約44点程であり、また、デコードの際を考慮しても、最終的に得たいサンプル数は最大でも63個程度であり、このような間引かれた演算を行うという性質が生かされていない。

【0013】本発明は、このような実情に鑑みてなされたものであり、演算量を低減しながらも、エンコードの際には可変個数のデータを一定個数に変換することができ、またデコードの際には一定個数のデータを可変個数のデータに変換することができるようなデータ数変換方法の提供を目的とする。

【0014】

【課題を解決するための手段】本発明に係るデータ数変換方法は、ブロック内の波形データ又は波形を表すパラメータデータの個数が可変とされたデータをブロック毎に一定の個数の基準データと比較するために上記可変個数の入力データを上記一定個数に変換する工程を含むデータ数変換方法において、上記可変個数のデータが入力される帯域制限型オーバーサンプリングのためのFIRフィルタで、上記入力データのサンプル点に対してそれぞれ異なる複数の位相に対応した複数の係数セットの内の上記一定個数のデータの各位置に対応する係数セットを用いることにより、出力として必要な上記一定個数のデータのみを求める工程を有することを特徴として上記課題を解決する。

【0015】また、他の発明に係るデータ数変換方法は、ブロック内の波形データ又は波形を表すパラメータデータの個数が可変とされたデータをブロック毎に一定の個数の基準データと比較するために上記可変個数の入力データを上記一定個数に変換する工程を含むデータ数変換方法において、上記可変個数のデータが入力される帯域制限型オーバーサンプリングのためのFIRフィルタで、上記入力データのサンプル点に対してそれぞれ異なる複数の位相に対応した複数の係数セットの内の上記一定個数のデータの各位置の近傍の位置に対応する係数セットを用いることにより、中間的な出力データを求める工程と、上記中間的な出力データを補間して最終出力に必要とされる一定個数のデータを求める工程とを有することを特徴として上記課題を解決する。

【0016】ここで、上記可変個数の入力データとしては、音声信号や音響信号等のオーディオ信号を、周波数軸上データに変換したものを用いることができ、具体的には、例えば、MBE（Multiband Excitation: マルチバンド励起）符号化の場合のスペクトル包絡振幅データや、SBE（Singleband Excitation: シングルバンド励

起) 符号化、ハーモニック (Harmonic) 符号化、SBC (Sub-band Coding: 帯域分割符号化)、LPC (Linear Predictive Coding: 線形予測符号化)、あるいはDCT (離散コサイン変換)、MDCT (モデファイドDCT)、FFT (高速フーリエ変換) 等におけるスペクトル振幅データやそのパラメータ (LSPパラメータ、 α パラメータ、 k パラメータ等) データ等を用いることができる。また、一定個数に変換されたデータを、ベクトル量子化することが考えられ、このベクトル量子化の前に、ブロック毎に一定個数のデータのブロック間の差分をとり、このブロック間差分データに対してベクトル量子化を施すようにしてもよい。

【0017】また、他の発明に係るデータ数変換方法は、ブロック内の波形データ又は波形を表すパラメータデータが一定個数とされたデータをブロック内のデータ内容に応じた個数のデータに変換するデータ数変換方法において、上記一定個数のデータが入力される帯域制限型オーバーサンプリングのためのFIRフィルタで、上記入力データのサンプル点に対してそれぞれ異なる複数の位相と対応した複数の係数セットの内の上記ブロック内のデータ内容に応じた個数のデータの各位置に対応する係数セットを用いることにより、出力として必要な上記ブロック内のデータ内容に応じた個数のデータのみを求める工程を有することを特徴として上記課題を解決する。

【0018】また、他の発明に係るデータ数変換方法は、ブロック内の波形データ又は波形を表すパラメータデータが一定個数とされたデータをブロック内のデータ内容に応じた個数のデータに変換するデータ数変換方法において、上記一定個数のデータが入力される帯域制限型オーバーサンプリングのためのFIRフィルタで、上記入力データのサンプル点に対してそれぞれ異なる複数の位相と対応した複数の係数セットの内の上記ブロック内のデータ内容に応じた個数のデータの各位置の近傍の位置に対応する係数セットを用いることにより、中間的な出力データを求める工程と、上記中間的な出力データを補間して最終出力に必要とされる上記ブロック内のデータ内容に応じた個数のデータを求める工程とを有することを特徴として上記課題を解決する。

【0019】上記中間的な出力データは、最終的に必要とされるデータの一個あたりそれぞれ少なくとも2個以上であればよい。

【0020】

【作用】一定個数の非線形圧縮データに変換されて基準データと非線形領域で比較することができ、ブロック間差分をとってベクトル量子化することが可能となる。また、データ数変換前のブロック内のデータ値の連続性を高め、ブロック端点でリングング等の生じない品質の高いデータ数変換が行える。

【0021】

【実施例】以下、本発明に係るデータ数変換方法の実施例について、図面を参照しながら説明する。図1は本発明の第1の実施例となるデータ数変換方法の概略構成を示している。

【0022】この第1の実施例は後述するMBEボコーダに適用される。すなわち、MBEボコーダにより算出されたスペクトルエンベロープの個数が可変とされた振幅データを一定個数に変換する方法である。

【0023】図1において、入力端子11には、後述するMBEボコーダにより算出されたスペクトルエンベロープの振幅データ等が供給されている。この振幅データは、例えば図2のAに示すようなスペクトルを有する音声信号を分析して、ピッチ周波数(角周波数) ω を求め、このピッチ周波数 ω に応じたスペクトルの周期性を考慮して、各高調波(ハーモニクス)位置での振幅から、図2のBに示すようなスペクトル包絡(エンベロープ)を表す振幅データとして求められる。この振幅データの個数は一定の有効帯域(例えば200~3400Hz)内でピッチ周波数 ω に依存して変化する。そこで、図2のCに示すように一定の固定周波数(角周波数) ω_c の各高調波位置での上記スペクトル包絡の振幅データを求めることで、データ個数を一定にできる。

【0024】図1の例では、入力端子11からの可変数M個(例えば $M=8\sim63$ である)の入力データを、非線形圧縮部12にて例えばdB領域に圧縮(対数圧縮)した後、データ個数変換本体部13にて一定個数のデータに変換している。データ個数変換本体部13は、スペクトルエンベロープ拡張部14、帯域制限型FIRフィルタ15及び直線補間部16から成っている。

【0025】入力されたブロック毎の可変数M個の入力データは、非線形圧縮部12で非線形圧縮され、スペクトルエンベロープ拡張部14でスペクトルエンベロープの両端の値を繰り返して前後に延長される。この両端が前後に延長されたスペクトルエンベロープは、FIRフィルタ15に供給される。このFIRフィルタ15は入力データのサンプル点に対してそれぞれ異なる複数の位相と対応した複数の係数セットの内の上記一定個数のデータの各位置の近傍の位置に対応する係数セットを用いることにより、中間的な出力データを求める。この中間的な出力データは、直線補間部16に供給され、直線補間されて最終出力に必要とされる一定個数のデータとなり、出力端子17から出力される。

【0026】ここで、後述するMBEボコーダにおいて算出されるM個($m_{MX}+1$ 個)の振幅データ列をa

[m]とする。mは上記高調波(ハーモニクス)の次数あるいはバンド番号であり、 m_{MX} が最大値であるが、 $m=0$ のバンドの振幅データも含めて、全バンドの振幅データの個数は $m_{MX}+1$ 個となる。この振幅データa

[m]を、非線形圧縮部12にて例えばdB領域に変換する。すなわち得られたデータを $a_{dB}[m]$ とすると

き、

$$a_{dB} [m] = 20 \log_{10} a [m] \quad \dots (1)$$

である。この対数変換された振幅データ $a_{dB} [m]$ の個数 $m_{MX} + 1$ は、上述したようにピッチに依存して変化するため、一定個数の振幅データ $b [m]$ に変換する。これは一種のサンプリングレート（サンプルレート）変換である。なお、非線形圧縮部 12 での圧縮処理は、dB 領域への対数圧縮の他に、例えばいわゆる μ -law や α -law のような疑似対数圧縮処理を施してもよい。このように、振幅を圧縮することにより、能率的な符号化が実現される。

【0027】MBE ボコーダに入力される時間軸上の音声信号に対するサンプリング周波数 f_s は、通常 8 kHz で、全帯域幅は 3.4 kHz（ただし有効帯域は 200 ~ 3400 Hz）であり、女声の高い方から男声の低い方までのピッチラグ（ピッチ周期に相当するサンプル数）は、20 ~ 147 程度である。従って、ピッチ（角）周波数 ω は、 $8000/147 \approx 54$ (Hz) から $8000/20 = 400$ (Hz) 程度までの間で変動することになる。従って、周波数軸上で上記 3.4 kHz までの間に約 8 ~ 63 本のピ

$$a_{JdB} [m] = \begin{cases} a_{dB} [0] & \text{for } -(f_0-1)/2 \leq m < 0 \\ a_{dB} [m] & \text{for } 0 \leq m < M \\ a_{dB} [M-1] & \text{for } M \leq m < M+(f_0-1)/2 \end{cases}$$

【0030】となる。ここで f_0 は例えば 9 で、次に使用する FIR フィルタの（オーバーサンプリング後のサンプリングレートでみた）次数 F_0 例えば 65 と、 $F_0 = O_S \times (f_0 - 1) + 1$ という関係にある定数である。また、 $f_0 - 1$ は、このスペクトルエンベロープ拡張を一種のオーバーサンプリングと考えたとき、オーバーサンプリングする前のサンプリングレートでみたときのフィルタ次数であり、 F_0 は、オーバーサンプリング後のサンプリングレートでみたときのフィルタの次数である。また、 O_S は、オーバーサンプリングの比率（レシオ）である。図 3 はこの $a_{JdB} [m]$ を示す図である。すなわち、この $a_{JdB} [m]$ は、 $0 \leq m < M$ の区間に示される元の波形 $a_{dB} [m]$ の左端 F_0 を $a_{dB} [0]$ のまま $-(f_0 - 1)$ まで延長し、右端部を最後のデータである $a_{dB} [M-1]$ のまま $M + (f_0 - 1)/2$ まで延長している。

【0031】本来、帯域制限型のオーバーサンプリングは、例えば $(O_S - 1)$ 個のデータを 0 詰めしたものに F_0 次のフィルタを通したものとしてもよいが、0 データに対する積和は無視してよい。そのため、帯域制限型のオーバーサンプリングは、8 つの位相の係数セット

($P = 0 \dots 7$) の各セット毎に $(f_0 - 1)$ 個の係数からなり、もとの F_0 個の係数をオーバーサンプリン

グしたパルス（ハーモニックス）が立つことになる。すなわち、周波数軸上の dB 領域の波形として、8 サンプル乃至 63 サンプルから成る $m_{MX} + 1$ 個のデータを、一定のサンプル数、例えば 44 サンプルに、サンプル数変換を行うわけである。これが、図 2 の C に示すように、一定のピッチ周波数（角周波数） ω_c 毎のハーモニックスの位置のサンプルを求めることに相当する。

【0028】次にスペクトルエンベロープ拡張部 14 は、上述したように非線形圧縮部 12 で非線形圧縮され、 $a_{dB} [m]$ の配列で表せる $m_{MX} + 1$ 個のスペクトルエンベロープの両端の値を前後に延長する。これはスペクトルエンベロープの端点におけるリングングの発生を防ぐために行われる。このようにしてできた数列を $a_{JdB} [m]$ とすると、この $a_{JdB} [m]$ は $-(f_0 - 1)/2 \leq m < M + (f_0 - 1)/2$ の範囲で、

【0029】

【数 1】

... (2)

グするフィルタ処理とみることができる。

【0032】図 4 は F_0 を 65、 f_0 を 9、 O_S を 8 とした場合の FIR フィルタ 15 の位相の係数を示す図である。図 4 の A は、 -4π から 4π までの位相の変化範囲で F_0 個（65 個）の係数の大きさを振幅値として示している。位相変化が 0π のとき係数値は 1 であるが、位相変化が $\pm 4\pi$ 、 $\pm 3\pi$ 、 $\pm 2\pi$ 、 $\pm \pi$ のときは 0 である。また、この図 4 の A は 0π の振幅を軸に左右対称となっている。図 4 の B は、 $P = 0 \dots 7$ の各位相の係数セットの持つ係数値が図 4 の A のどこにあたるかを示している。なお、この係数値は、周知の方法によって導出できる。

【0033】ここで、本発明においては、上記入力データのサンプル点に対してそれぞれ異なる複数の位相と対応した複数の係数セットの内の上記一定個数のデータの各位置に対応する係数セットを用いることにより、出力として必要なデータそのもの、あるいは必要とされるデータの近傍のデータを求めることにより、演算自体を間引いて演算量を減らすものである。

【0034】図 5 は、図 4 に示された係数セット ($P = 0 \dots 7$) を用いて $a_{JdB} [m]$ をフィルタリングし、出力として必要とされる一定個数のデータ $b [m]$ の内の任意の 1 個を得るための演算を説明するための図

である。

【0035】図5のAは、 a_{dB} [m] を示す。この a_{dB} [m] から b [m] を得るには、上記スペクトルエンベロープ拡張部14で a_{dB} [m] の両端を延長して図3に示したような a_{JdB} [m] を先ず得る。 i は可変数 M 個のデータのインデックスである。

【0036】例えば、図5のAに示された b 点における b [m] を求めようとする場合を以下に述べる。この b 点に最も近傍の位置にある係数値のセットは $P=2$ の係数セットである。この $P=2$ の係数セットは図5のBに示すような各係数値を持っている。この各係数値を p_{20} 、 p_{21} 、 p_{22} 、 p_{23} 、 p_{24} 、 p_{25} 、 p_{27} とする。すると、 b 点の b [m] は、インデックス $i=0$ のデータと係数値 p_{20} の乗算値と、インデックス $i=1$ のデータ

と係数値 p_{21} の乗算値と、インデックス $i=2$ のデータと係数値 p_{22} の乗算値と、インデックス $i=3$ のデータと係数値 p_{23} の乗算値と、インデックス $i=4$ のデータと係数値 p_{24} の乗算値と、インデックス $i=5$ のデータと係数値 p_{25} の乗算値と、インデックス $i=6$ のデータと係数値 p_{26} の乗算値と、インデックス $i=7$ のデータと係数値 p_{27} の乗算値との合計8個の乗算値の和として表せる。

【0037】今、 F_0 個の係数を $0 \leq k < F_0$ の範囲で $\text{coef}[k]$ とすると $0 \leq m < M \cdot O_s$ の範囲での b [m] は、次の(3)式で示される。

【0038】

【数2】

$$b(i \times O_s + P) = \sum_{j=1}^{f_0-1} \text{coef}[j \times O_s - P] a_{JdB}[i + j - (f_0-1)/2]$$

$$\text{for} \begin{cases} 0 \leq i < M \\ 0 \leq P < O_s \end{cases}$$

... (3)

【0039】この(3)式より、例えば、上記 b 点の b [m] は、 $O_s=8$ であれば、 $b[3 \times 8 + 2]$ であり、 $b[26]$ となり、 $b[26]$ のサンプルデータの振幅値を求めることになる。

【0040】ここで、 $O_s=8$ 、 $f_0=9$ とすると上記(3)式は、

【数3】

$$b(8 \times i + P) = \sum_{j=1}^8 \text{coef}[8 \times j - P] a_{JdB}[i + j - 4]$$

$$\text{for} \begin{cases} 0 \leq i < M \\ 0 \leq P < O_s \end{cases}$$

... (4)

となる。

【数4】

【0041】ここで、 $P=0$ のとき上記(4)式は、

$$b(8 \times i) = \sum_{j=1}^8 \text{coef}[8 \times j] a_{JdB}[i + j - 4]$$

$$= a_{JdB}[i]$$

$$\text{for} \quad 0 \leq i < M$$

... (5)

となり、 $i=0, 1 \dots 7$ のデータの振幅をそのまま求めることになる。

は、

【数5】

【0042】また、 $P=1 \dots 7$ のとき上記(4)式

$$b[8 \times i + P] = \sum_{j=1}^8 \text{coef}[8 \times j - P] a_{\text{JdB}}[i + j - 4]$$

for $0 \leq i < M$

... (6)

【0043】となり、 $a_{\text{JdB}}[i-3]$ 、 $a_{\text{JdB}}[i-2]$ 、 $a_{\text{JdB}}[i-1]$ 、 $a_{\text{JdB}}[i-0]$ 、 $a_{\text{JdB}}[i+1]$ 、 $a_{\text{JdB}}[i+2]$ 、 $a_{\text{JdB}}[i+3]$ 、 $a_{\text{JdB}}[i+4]$ の8個のデータに対し、coef[8-P]、coef[16-P]、coef[24-P]、coef[32-P]、coef[40-P]、coef[48-P]、coef[56-P]、coef[64-P]の8個の係数が各々乗算され、その8個の乗算値が全て加算されてb[m]が得られることが分かる。

【0044】例えば、上記図5のAに示されたb点のb[m]の例では、 $i=3$ 、 $P=2$ であるので、 $a_{\text{JdB}}[0]$ 、 $a_{\text{JdB}}[1]$ 、...、 $a_{\text{JdB}}[7]$ の8個のデータに対し、coef[6]、coef[14]、...、coef[62]の8個の係数が各々乗算され、その8個の乗算値が全て加算されてb[26]が得られる。

【0045】また、例えば、 $i=0$ 、 $P=3$ のb[m]を求める場合は、上記スペクトルエンベロープ拡張部14で得たデータ $a_{\text{JdB}}[-3]$ 、 $a_{\text{JdB}}[-2]$ 、 $a_{\text{JdB}}[-1]$ の計3個のデータにそれぞれcoef[5]、coef[11]、coef[22]の計3個の係数を乗算した3個の乗算値と、 $a_{\text{JdB}}[0]$ 、 $a_{\text{JdB}}[1]$ 、 $a_{\text{JdB}}[2]$ 、 $a_{\text{JdB}}[3]$ 、 $a_{\text{JdB}}[4]$ の計5個のデータにそれぞれcoef[29]、coef[37]、coef[45]、coef[53]、coef[61]の5個の係数を乗算した5個の乗算値とからなる計8個の乗算値が加算されてb[3]が得られる。

【0046】ここで、 $a_{\text{JdB}}[i+j-(f_0-1)/2]$ の〔〕内が $i_{\min}-(f_0-1)/2=1-(f_0-1)/2$ 、 $i_{\max}+(f_0-1)/2=m_{\text{MX}}+(f_0-1)/2$ に関しては、スペクトルエンベロープ拡張部14により、データが拡張されているので問題はない。ここで、b[m]の一点を求めるのに必要な積は (f_0-1) 回である。

【0047】ところで、以上の説明においては、(例えば8倍の)オーバーサンプル点のいずれかの位置に上記最終的に必要とされる一定個数(例えば44個)のデータの位置が一致するものとして説明したが、現実には、このような一致を得るためにはオーバーサンプルの比率

(倍数)を極めて高くとることが必要とされ、フィルタ係数の個数が膨大なものとなることより、最終的には必要とされるデータの位置の近傍(例えば前後の2点)のオーバーサンプル点のデータを中間的な出力として上記フィルタリング演算により求め、この中間的な出力を補間処理することで、上記最終的に必要とされるデータを求めることが好ましい。

【0048】すなわち、上記FIRフィルタ15からのFIR出力は、直線補間部16に供給される。この直線補間部16は、上記FIRフィルタ15からの少なくとも2つのFIR出力を直線補間し、必要な出力点を得る。例えば、図6において点 A_0 を直線補間で求めるには、その点 A_0 を挟む2点 A_{-1} 、 A_1 がFIRフィルタで算出されていればよい。したがって、データ個数変換本体部13で求められる最終的なエンコーダでの出力点の個数を44点とすれば、 $44 \times 2 (=88)$ 点が上記FIRフィルタ15で算出されればよい。

【0049】この必要な出力点を得るための2点 A_{-1} 、 A_1 のb[m]を求める処理について図7のフローチャートを用いて説明する。ステップS1では、入力側の角周波数を O_S (サンプリングレシオ)で割った ω_{0f} と、出力側の角周波数 ω_0 を求める。第1の実施例では、 O_S (例えば8)倍のサンプリングを行っているので、スペクトルは、入力側の角周波数の $1/O_S$ のインターバルで立っている。そのため、 O_S で割った値 ω_{0f} を出す。 $0 \sim \pi$ までを例えば1024/M×1/ O_S となる。また、欲しい点(出力側)の角周波数は ω_0 であり、この ω_0 が $1024/M'$ となる。ここで、 M' は、出力側のハーモニクス数である。

【0050】ステップS2では、入力側ハーモニクスのインデックスi及び出力側ハーモニクスのインデックスiiを初期化する。ステップS3では、上記係数セットPを初期化する。

【0051】ステップS4では、入力側ハーモニクスのインデックスiと係数セットPとにより求めたいデータの位置 A_0 を検索(スキャン)する。すなわち、求めたいデータの位置 $A_0 (= \omega_0 \times ii)$ を、iとPによるス

キャンの位置 $A_1 (= i \times O_s + P + 1)$ が越えたか否かを判定する。例えば、始めは上記ステップS2、S3で i と P が初期化されているので $i=0$ 、 $P=0$ として検索する。ここで、YESを判定するとステップS5に進み、NOを判定するとステップS7に進む。

【0052】ステップS5では、求めたいデータの位置 $A_0 (= \omega_0 \times ii)$ を越えた i と P によるスキャンの位置 $A_1 (= i \times O_s + P + 1)$ での $b[m]$ 、すなわち、 $b[i \times O_s + P + 1]$ とその一つ前(A_{-1})の $b[i \times O_s + P]$ とを求める。この $b[i \times O_s + P + 1]$ と $b[i \times O_s + P]$ とは上記求めたいデータの位置 $A_0 (= \omega_0 \times ii)$ を挟み込むような位置(A_1 と A_{-1} の間)での $b[m]$ となる。

【0053】ステップS6では、次に求めたいデータの位置を移動するため、出力側ハーモニクスのインデックス ii をインクリメントする。ステップS7では、スキャンの位置を移動するために係数セット P をインクリメントする。このとき i は0のままである。すなわち、 $i=0$ のまま、 P を0から1に変える。

【0054】ステップS8では、係数セット P が O_s の値と一致したか否かを判定する。 P は0・・・7までの8個であり、 O_s も8としている。ここで、YESを判定するとステップS9に進み、NOを判定するとステップS4に進む。

【0055】ステップS9では、入力側ハーモニクスのインデックス i をインクリメントする。そして、ステップS10に進む。

【0056】ステップS10では、上記 i が可変個数のデータと数(M 個)と等しくなったか否かを判定する。ここでYESを判定するとこのフローは終了となり、NOを判定するとステップS3に戻る。

【0057】以上のフローチャートより、本実施例は O_s (ここでは $O_s=8$)倍でオーバーサンプリングピッチ(角周波数) ω_{0f} のインターバルで周波数をインクリメントしてゆき出力として欲しい点を越えたところでの $b[m]$ とその一つ手前の $b[m]$ とを求めている。このようにすれば、出力点を直線補間で求めるのに必要な左右の点が全て算出されることになる。

【0058】次に、上記図7に示したフローチャートの処理によって求められた、出力として欲しい点を越えたところでの $b[m]$ とその一つ手前の $b[m]$ を直線補間部16により直線する処理を図8のフローチャートを用いて説明する。

【0059】ステップS21では、出力角周波数 ω_0 と、入力角周波数 ω_{0f} とを求める。これは、上記図7に示したステップS1と同様である。ステップS22では、以後のフローが入力側のハーモニクスの8倍のインデックス i でインクリメントされるのでこの i を初期化する。ステップS23では、 $i=0$ になっているか否かを判別する。ここでYESを判別するとステップS24

に進み、NOを判別するとステップS25に進む。

【0060】ステップS24、ステップS25では、図6に示すようにある一つの区間に着目して、その幅を b_w とし、上限を u_b 、下限を l_b としている。この上限 u_b は、 $\text{inint}(i+1) \times \omega_{0f}$ となり、下限 l_b は、 $\text{inint } i \times \omega_{0f}$ となる。ここで、 inint は $\text{inint}(x)$ とすると、 x に最も近い数を返す関数である。また、上記下限 l_b は、一回前の上限 u_b となる。したがって、 b_w は、上限 u_b と下限 l_b との差になる。

【0061】上記ステップS24では、下限 l_b を0とし、ステップS26に進む。上記ステップS25では、下限 l_b と上限 u_b とを一致させる。

【0062】ステップS26では、上述したように上限 u_b を $\text{inint}(i+1) \times \omega_{0f}$ と設定する。ステップS27では、上限 u_b と下限 l_b との差である b_w を求める。そして、この b_w の間をスキャンして、直線補間値 $c[ii]$ を求める。

【0063】ステップS28では、図6に示す求めようとする $c[ii]$ と下限 l_b との差 i_{dx} を0に設定する。すなわち、 $i_{dx}=0$ の位置(下限 l_b と一致)からスキャンを開始するスキャン開始位置を設定する。

【0064】ステップS29では、上述したように下限 l_b からスキャン j を開始する。ステップS30では、スキャン j が求めようとする $c[ii]$ の位置と一致したか否かを判別する。ここで、YESを判別するとステップS31に進み、NOを判別するとステップS32に進む。

【0065】ステップS31では、位置関係に関連する重み付けを考慮した $c[ii]$ を求める。ここで、例えば、 i_{dx} が0のときは、 $c[ii]=b[i]$ となり、 i_{dx} が b_w のときは、 $c[ii]=b[i+1]$ となる。

【0066】ステップS32では、 i_{dx} をインクリメントする。そして、ステップS33では、出力ハーモニクスのインデックス ii が出力ハーモニクスの数 M' より大きくなったか否かを判別する。ここで、YESを判別すると、このフローは終了となり、NOを判別するとステップS34に進む。

【0067】ステップS34では、スキャン j の繰返しを始める。ステップS35では、スキャン j が上限 u_b まで到達したか否かを判別する。ここで、YESを判別するとステップS36に進み、NOを判別するとステップS30に戻る。

【0068】ステップS36では、入力側のハーモニクス i をインクリメントする。ステップS37では、 i が入力ハーモニクス M と O_s との積よりも大きくなったか否かを判別する。ここで、YESを判別するとこのフローは終了となるが、NOを判別するとステップS23に戻る。

【0069】以上のフローチャートより、本実施例は、上記図7のフローチャートの処理で求めた $b[m]$ を直

線補間部16により直線補間するだけで、必要な点だけを求められる。

【0070】このように第1の実施例は、必要な点のみを求めることによって、個数が可変とされたデータを一定個数にすることができる。そのため、演算量が減少する。

【0071】このようにして、一定サンプル数のデータに変換した数列に必要に応じてブロック間、あるいはフレーム間で差分をとり、ベクトル量子化を施して、そのインデックスを伝送するようにすればよい。

【0072】上述した第1の実施例は、MBEボコーダにより算出されたスペクトルエンベロープの個数が可変とされた振幅データを一定個数に変換する方法であったが、以下、第2の実施例として、一定個数にされたデータをデータ内容に応じた個数のデータに変換するデータ個数変換方法を説明する。この第2の実施例は例えば音声信号を合成するデコーダ側に適用される。すなわち、デコーダ側では、上記インデックスより、ベクトル量子化及び逆量子化された数列の一定個数とされた波形データを得て、そのデータ列を、同様の方法で、すなわち帯域制限オーバーサンプリング、直線補間等を施すことにより、データの内容に応じた個数のM個の数列に変換する。

【0073】図9は第2の実施例の概略構成を示している。上記第1の実施例において、一定個数とされた入力データは入力端子21を介してデータ個数変換本体部22に供給され、このデータ個数変換本体部22で可変個数のデータとされて出力端子26から出力される。このデータ個数変換本体部22は、スペクトルエンベロープ拡張部23、帯域制限型FIRフィルタ24及び直線補間部25から成っている。

【0074】入力されたブロック毎に一定個数の入力データは、スペクトルエンベロープ拡張部23でスペクトルエンベロープの両端の値を延長される。この両端が前後に延長されたスペクトルエンベロープは、FIRフィルタ24に供給される。このFIRフィルタ24はスペクトルエンベロープが延長されることによりデータ個数が拡大されたデータのサンプル点に対しそれぞれ異なる複数の位相と対応した複数の係数セットの内の一定個数のデータの各位置の近傍の位置に対応する係数セットを用いることにより、中間的な出力データを求める。そして、この中間的な出力データは直線補間部25に供給される。この直線補間部25は上記中間的な出力データを直線補間し、出力端子26から間引きされ、データ内容に応じた可変個数のデータを出力する。

【0075】この第2の実施例は、必要な点のみを求めることによって、個数が一定とされたデータをデータ内容に応じた個数に変換することができる。そのため、演算量が減少される。

【0076】ここで、第1の実施例による乗算の回数

は、求めるデータの個数を44個とすれば、その2倍の88個のデータに対し、8回の乗算が施されることになり、1024回の乗算となる。これは、上述した高速化手法を用いたFFT、IFFTの乗算の回数の合計31488回の $1/45$ となる。また、第2の実施例による乗算の回数は、求めるデータの個数を60個とすれば、その2倍の120個のデータに対し、8回の乗算がほどこされることになる。これは、上述した高速化手法を用いたFFT、IFFTの乗算の回数の合計31488回の $1/30$ となる。

【0077】次に、上述したようなデータ数変換方法が適用可能な、音声信号の合成分析符号化装置（いわゆるボコーダ）の一種のMBE（Multiband Excitation: マルチバンド励起）ボコーダの具体例について、図面を参照しながら説明する。以下に説明するMBEボコーダは、D.W. Griffin and J.S. Lim, "Multiband Excitation Vocoder," IEEE Trans. Acoustics, Speech, and Signal Processing, vol. 36, No. 8, pp. 1223-1235, Aug. 1988に開示されているものであり、従来のPARCOR（Partial auto-CORrelation: 偏自己相関）ボコーダ等では、音声のモデル化の際に有声音区間と無声音区間とをブロックあるいはフレーム毎に切り換えていたのに対し、MBEボコーダでは、同時刻（同じブロックあるいはフレーム内）の周波数軸領域に有声音（Voiced）区間と無声音（Unvoiced）区間とが存在するという仮定でモデル化している。

【0078】図10は、上記MBEボコーダに本発明を適用した実施例の全体の概略構成を示すブロック図である。この図10において、入力端子101には音声信号が供給されるようになっており、この入力音声信号は、HPF（ハイパスフィルタ）等のフィルタ102に送られて、いわゆるDC（直流）オフセット分の除去や帯域制限（例えば200～3400Hzに制限）のための少なくとも低域成分（200Hz以下）の除去が行われる。このフィルタ102を介して得られた信号は、ピッチ抽出部103及び窓かけ処理部104にそれぞれ送られる。ピッチ抽出部103では、入力音声信号データが所定サンプル数N（例えばN=256）単位でブロック分割され（あるいは方形窓による切り出しが行われ）、このブロック内の音声信号についてのピッチ抽出が行われる。このような切り出しブロック（256サンプル）を、例えば図11のAに示すようにLサンプル（例えばL=160）のフレーム間隔で時間軸方向に移動させており、各ブロック間のオーバーラップはN-Lサンプル（例えば96サンプル）となっている。また、窓かけ処理部104では、1ブロックNサンプルに対して所定の窓関数、例えばハミング窓をかけ、この窓かけブロックを1フレームLサンプルの間隔で時間軸方向に順次移動させている。

【0079】このような窓かけ処理を数式で表すと、

$$x_w(k, q) = x(q) w(kL - q) \quad \dots (7)$$

【0080】となる。この(7)式において、 k はブロック番号を、 q はデータの時間インデックス(サンプル番号)を表し、処理前の入力信号の q 番目のデータ $x(q)$ に対して第 k ブロックの窓(ウィンドウ)関数 $w(kL - q)$ により窓かけ処理されることによりデータ $x_w(k, q)$

q が得られることを示している。ピッチ抽出部103内での図11のAに示すような方形窓の場合の窓関数 $w_r(r)$ は、

$$[0081]$$

$$w_r(r) = 1 \quad 0 \leq r < N \quad \dots (8)$$

$$= 0 \quad r < 0, N \leq r$$

【0082】また、窓かけ処理部104での図11のB

に示すようなハミング窓の場合の窓関数 $w_h(r)$ は、

$$w_h(r) = 0.54 - 0.46 \cos(2\pi r / (N-1)) \quad \dots (9)$$

$$0 \leq r < N$$

$$= 0 \quad r < 0, N \leq r$$

【0083】である。このような窓関数 $w_r(r)$ あるいは $w_h(r)$ を用いるときの上記(7)式の窓関数 $w(r) (=w(kL - q))$ の非零区間は、

$$0 \leq kL - q < N$$

これを变形して、

$$kL - N < q \leq kL$$

【0084】従って例えば上記方形窓の場合に窓関数 $w_r(kL - q) = 1$ となるのは、図12に示すように、 $kL - N < q \leq kL$ のときとなる。また、上記(7)～(9)式は、長さ $N (=256)$ サンプルの窓が、 $L (=160)$ サンプルずつ前進してゆくことを示している。以下、上記(8)式、(9)式の各窓関数で切り出された各 N 点($0 \leq r < N$)の非零サンプル列を、それぞれ $x_{wr}(k, r)$ 、 $x_{wh}(k, r)$ と表すことにする。

【0085】窓かけ処理部104では、図13に示すように、上記(9)式のハミング窓がかけられた1ブロック256サンプルのサンプル列 $x_{wh}(k, r)$ に対して1792サンプル分の0データが付加されて(いわゆる0詰めされて)2048サンプルとされ、この2048サンプルの時間軸データ列に対して、直交変換部105により例えばFFT(高速フーリエ変換)等の直交変換処理が施される。

【0086】ピッチ抽出部103では、上記 $x_{wr}(k, r)$ のサンプル列(1ブロック N サンプル)に基づいてピッチ抽出が行われる。このピッチ抽出法には、時間波形の周期性や、スペクトルの周期的周波数構造や、自己相関関数を用いるもの等が知られているが、本実施例では、センタクリップ波形の自己相関法を採用している。このときのブロック内でのセンタクリップレベルについては、1ブロックにつき1つのクリップレベルを設定してもよいが、ブロックを細分割した各部(各サブブロック)の信号のピークレベル等を検出し、これらの各サブブロックのピークレベル等の差が大きいときに、ブロック内でクリップレベルを段階的にあるいは連続的に変化

$$S(j) = H(j) | E(j) |$$

と表現するようなモデルを想定している。ここで、 J は $\pi \omega_s = f_s / 2$ に対応し、サンプリング周波数 $f_s = 2\pi \omega_s$ が例えば8kHzのときには4kHzに対応する。

させるようにしている。このセンタクリップ波形の自己相関データのピーク位置に基づいてピーク周期を決めている。このとき、現在フレームに属する自己相関データ(自己相関は1ブロック N サンプルのデータを対象として求められる)から複数のピークを求めておき、これらの複数のピークの内の最大ピークが所定の閾値以上のときには該最大ピーク位置をピッチ周期とし、それ以外のときには、現在フレーム以外のフレーム、例えば前後のフレームで求められたピッチに対して所定の関係を満たすピッチ範囲内、例えば前フレームのピッチを中心として $\pm 20\%$ の範囲内にあるピークを求め、このピーク位置に基づいて現在フレームのピッチを決定するようにしている。このピッチ抽出部103ではオープンループによる比較的ラフなピッチのサーチが行われ、抽出されたピッチデータは高精度(ファイン)ピッチサーチ部106に送られて、クローズドループによる高精度のピッチサーチ(ピッチのファインサーチ)が行われる。

【0087】高精度(ファイン)ピッチサーチ部106には、ピッチ抽出部103で抽出された整数(インテジャー)値の粗(ラフ)ピッチデータと、直交変換部105により例えばFFTされた周波数軸上のデータとが供給されている。この高精度ピッチサーチ部106では、上記粗ピッチデータ値を中心に、 $0.2 \sim 0.5$ きざみで±数サンプルずつ振って、最適な小数点付き(フローティング)のファインピッチデータの値へ追い込む。このときのファインサーチの手法として、いわゆる合成による分析(Analysis by Synthesis)法を用い、合成されたパワースペクトルが原音のパワースペクトルに最も近くなるようにピッチを選んでいる。

【0088】このピッチのファインサーチについて説明する。先ず、上記MBEボコーダにおいては、上記FFT等により直交変換された周波数軸上のスペクトルデータとしての $S(j)$ を

$$0 < j < J \quad \dots (10)$$

上記(10)式中において、周波数軸上のスペクトルデータ $S(j)$ が図14のAに示すような波形のとき、 $H(j)$ は、図14のBに示すような元のスペクトルデータ S

(j) のスペクトル包絡線 (エンベロープ) を示し、E(j) は、図14のCに示すような等レベルで周期的な励起信号 (エキサイテーション) のスペクトルを示している。すなわち、FFTスペクトルS(j) は、スペクトルエンベロープH(j) と励起信号のパワースペクトル|E(j)| との積としてモデル化される。

【0089】上記励起信号のパワースペクトル|E(j)| は、上記ピッチに応じて決定される周波数軸上の波形の周期性 (ピッチ構造) を考慮して、1つの帯域 (バンド) の波形に相当するスペクトル波形を周波数軸上の各バンド毎に繰り返すように配列することにより形成される。この1バンド分の波形は、例えば上記図13に示すような256サンプルのハミング窓関数に1792サン

$$\varepsilon_m = \sum_{j=a_m}^{b_m} \{ |S(j)| - |A_m| |E(j)| \} \quad \dots (11)$$

で表せる。このエラー ε_m を最小化するような $|A_m|$ は、

$$\frac{\partial \varepsilon_m}{\partial |A_m|} = 2 \sum_{j=a_m}^{b_m} \{ |S(j)| - |A_m| |E(j)| \} |E(j)|$$

$$\therefore |A_m| = 2 \sum_{j=a_m}^{b_m} |S(j)| |E(j)| / \sum_{j=a_m}^{b_m} |E(j)|^2 \quad \dots (12)$$

【0093】となり、この(12)式の $|A_m|$ のとき、エラー ε_m を最小化する。このような振幅 $|A_m|$ を各バンド毎に求め、得られた各振幅 $|A_m|$ を用いて上記(11)式で定義された各バンド毎のエラー ε_m を求める。次に、このような各バンド毎のエラー ε_m の全バンドの総和値 $\sum \varepsilon_m$ を求める。さらに、このような全バンドのエラー総和値 $\sum \varepsilon_m$ を、いくつかの微小に異なるピッチについて求め、エラー総和値 $\sum \varepsilon_m$ が最小となるようなピッチを求める。

【0094】すなわち、上記ピッチ抽出部103で求められたラフピッチを中心として、例えば0.25きざみで上下に数種類ずつ用意する。これらの複数種類の微小に異なるピッチの各ピッチに対してそれぞれ上記エラー総和値 $\sum \varepsilon_m$ を求める。この場合、ピッチが定まるとバンド幅が決まり、上記(13)式より、周波数軸上データのパワースペクトル|S(j)| と励起信号スペクトル|E(j)| とを用いて上記(11)式のエラー ε_m を求め、その全バンドの総和値 $\sum \varepsilon_m$ を求めることができる。このエラー総和値 $\sum \varepsilon_m$ を各ピッチ毎に求め、最小となるエラー総和値に対応するピッチを最適のピッチとして決定

するわけである。以上のようにして高精度ピッチサーチ部106で最適のファイン (例えば0.25きざみ) ピッチが求められ、この最適ピッチに対応する振幅 $|A_m|$ が決定される。

【0090】次に、上記ピッチに応じて分割された各バンド毎に、上記H(j) を代表させるような (各バンド毎のエラーを最小化するような) 値 (一種の振幅) $|A_m|$ を求める。ここで、例えば第mバンド (第m高調波の帯域) の下限、上限の点をそれぞれ a_m 、 b_m とするとき、この第mバンドのエラー ε_m は、

【0091】

【数6】

【0092】

【数7】

【0095】以上ピッチのファインサーチの説明においては、説明を簡略化するために、全バンドが有声音 (Voiced) の場合を想定しているが、上述したようにMBEポコードにおいては、同時刻の周波数軸上に無声音 (Unvoiced) 領域が存在するというモデルを採用していることから、上記各バンド毎に有声音/無声音の判別を行うことが必要とされる。

【0096】上記高精度ピッチサーチ部106からの最適ピッチ及び振幅 $|A_m|$ のデータは、有声音/無声音判別部107に送られ、上記各バンド毎に有声音/無声音の判別が行われる。この判別のために、NSR (ノイズ to シグナル比) を利用する。すなわち、第mバンドのNSRは、

【0097】

【数8】

$$NSR = \frac{\sum_{j=a_m}^{b_m} \{ |S(j)| - |A_m| |E(j)| \}^2}{\sum_{j=a_m}^{b_m} |S(j)|^2} \quad \dots (13)$$

【0098】と表せ、このNSR値が所定の閾値（例えば0.3）より大のとき（エラーが大きい）ときには、そのバンドでの $|A_m| |E(j)|$ による $|S(j)|$ の近似が良くない（上記励起信号 $|E(j)|$ が基底として不適当である）と判断でき、当該バンドをUV（Unvoiced、無声音）と判別する。これ以外のときは、近似がある程度良好に行われていると判断でき、そのバンドをV（Voiced、有声音）と判別する。

【0099】次に、振幅再評価部108には、直交変換部105からの周波数軸上データ、高精度ピッチサーチ

部106からのファインピッチと評価された振幅 $|A_m|$ との各データ、及び上記有声音／無声音判別部107からのV／UV（有声音／無声音）判別データが供給されている。この振幅再評価部108では、有声音／無声音判別部107において無声音（UV）と判別されたバンドに関して、再度振幅を求めている。このUVのバンドについての振幅 $|A_m|_{UV}$ は、

【0100】

【数9】

$$|A_m|_{UV} = \sqrt{\frac{\sum_{j=a_m}^{b_m} |S(j)|^2}{(b_m - a_m + 1)}} \quad \dots (14)$$

にて求められる。

【0101】この振幅再評価部108からのデータは、データ数変換（一種のサンプリングレート変換）部109に送られる。このデータ数変換部109は、上記ピッチに応じて周波数軸上での分割帯域数が異なり、データ数（特に振幅データの数）が異なることを考慮して、一定の個数にするためのものである。すなわち、例えば有効帯域を3400 kHzまでとすると、この有効帯域が上記ピッチに応じて、8バンド～63バンドに分割されることになり、これらの各バンド毎に得られる上記振幅 $|A_m|$ （UVバンドの振幅 $|A_m|_{UV}$ も含む）データの個数 $m_{MX}+1$ も8～63と変化することになる。このためデータ数変換部109では、この可変個数 $m_{MX}+1$ の振幅データを一定個数（例えば44個）のデータに変換している。

【0102】ここで本第1の実施例においては、上記図1～図8と共に説明したように、周波数軸上の有効帯域1ブロック分の振幅データに対して、ブロック内の両端のデータを延長してデータ個数を拡大し、帯域制限型FIRフィルタによるフィルタ処理を施し、さらに直線補間を施すことにより一定個数（例えば44個）のデータを得ている。

【0103】このデータ数変換部109からのデータ（上記一定個数の振幅データ）がベクトル量子化部110に送られて、所定個数のデータ毎にまとめられてベクトルとされ、ベクトル量子化が施される。ベクトル量子化部110からの量子化出力データは、CRC&レート1/2畳込み符号付加部111に供給されと共にフレームインターリーブ部112に供給される。また、上記高精度のピッチサーチ部106からの高精度（ファイン）

ピッチデータ及び上記有声音／無声音判別部107からの有声音／無声音（V／UV）判別データも上記CRC&レート1/2畳込み符号付加部111に供給される。

【0104】ここで、上記CRC&レート1/2畳込み符号付加部111は、上記ファインピッチデータ、V／UV判別データ及び量子化出力データを用いて、スペクトルエンベロープの量子化を階層的な構造とし、その出力インデックスの重要度を分けることで効果的に畳込み符号による誤り訂正を行う。

【0105】これは、本件出願人が特願平4-91422号において、提案した高能率符号化方法、すなわち、M次元ベクトルを、S次元（ $S \leq M$ ）ベクトルに次元低下させてベクトル量子化するような、階層構造化されたコードブックを有する量子化を行わせる方法と同様に誤り訂正符号の効果的な適用が可能となる方法である。

【0106】具体的に、このデコーダ側のビタビ符号&CRC検出は、以下のような原理である。図15は、ビタビ復号&CRC検出の原理を説明するための機能ブロック図である。例えば、音声符号器121から出力された音声パラメータのうち、聴覚上特に重要な部分（クラス1）80ビットとそれ以外の部分（クラス2）40ビットとに分ける。クラス1のうちさらに重要な50ビットについてCRC計算ブロック122によりCRCを計算し、7ビットの結果を得る。クラス1の80ビットとCRCの7ビットと畳込み符号化器の初期値を0に戻すためのテールビット5ビットの合計92ビットを畳込み符号化部123に入力し、184ビットの出力を得る。畳込み符号化された184ビットとクラス2ビットの40ビットの計224ビットにつき、2スロットインターリーブ器124により、インターリーブを行い、その出

力として224ビットを伝送する。

【0107】この2スロットインターリーブ器124に相当するのが図10のフレームインターリーブ部112であり、その出力が出力端子113から伝送される。

【0108】なお、これらの各データは、上記Nサンプル（例えば256サンプル）のブロック内のデータに対して処理を施すことにより得られるものであるが、ブロックは時間軸上を上記Lサンプルのフレームを単位として前進することから、伝送するデータは上記フレーム単位で得られる。すなわち、上記フレーム周期でピッチデータ、V/U判別データ、振幅データが更新されることになる。

【0109】次に、伝送されて得られた上記出力データに基づき音声信号を合成するための合成側（デコード側）の概略構成について、図16を参照しながら説明する。この図15において、入力端子131には、伝送されたきたCRC&レート1/2畳込み符号が付加された出力データが供給される。入力端子131からの出力データは、フレームデインターリーブ132に供給され、デインターリーブされる。デインターリーブされたデータは、ビタビ復号&CRC検出部133に供給され、復号化される。

【0110】そして、マスク処理部134が、フレームデインターリーブ132からのデータをマスク処理し、量子化振幅データを逆ベクトル量子化部135に供給する。

【0111】この逆量子化部135も階層構造化されており、各階層のインデックスデータに基づいて逆ベクトル化されたデータを合成して出力する。この逆量子化部135からの出力データは、データ数逆変換部136に送られて逆変換される。このデータ数逆変換部136では、上述した図9の説明と同様な（逆）変換が行われ、得られた振幅データが有声音合成部137及び無声音合

$$V_m(n) = A_m(n) \cos(\theta_m(n)) \quad 0 \leq n < L \quad \dots (15)$$

と表すことができる。全バンドの内のV（有声音）と判別された全てのバンドの有声音を加算（ $\Sigma V_m(n)$ ）して最終的な有声音V(n)を合成する。

【0114】この(15)式中の $A_m(n)$ は、上記合成フレームの先端から終端までの間で補間された第m高調波の振幅である。最も簡単には、フレーム単位で更新され

$$A_m(n) = (L-n)A_{0m}/L + nA_{Lm}/L \quad \dots (16)$$

の式により $A_m(n)$ を計算すればよい。

$$\theta_m(0) = m\omega_{01}n + n^2 m(\omega_{L1} - \omega_{01}) / 2L + \phi_{0m} + \Delta\omega n \quad \dots (17)$$

により求めることができる。この(17)式中で、 ϕ_{0m} は上記合成フレームの先端（ $n=0$ ）での第m高調波の位相（フレーム初期位相）を示し、 ω_{01} は合成フレーム先端（ $n=0$ ）での基本角周波数、 ω_{L1} は該合成フレームの終端（ $n=L$ ：次の合成フレーム先端）での基本角周波数をそれぞれ示している。上記(17)式中の $\Delta\omega$ は、

成部138に送られる。また、上記マスク処理部134は、符号化ピッチデータをピッチ復号化部139に供給する。このピッチ復号化部139で復号されたピッチデータは、データ数逆変換部136、有声音合成部137及び無声音合成部138に送られる。また、上記マスク処理部134は、V/U判別データを有声音合成部137及び無声音合成部138に供給する。

【0112】有声音合成部137では例えば余弦(cosine)波合成により時間軸上の有声音波形を合成し、無声音合成部138では例えばホワイトノイズをバンドパスフィルタでフィルタリングして時間軸上の無声音波形を合成し、これらの各有声音合成波形と無声音合成波形とを加算部140で加算合成して、出力端子141より取り出すようにしている。この場合、上記振幅データ、ピッチデータ及びV/U判別データは、上記分析時の1フレーム（Lサンプル、例えば160サンプル）毎に更新されて与えられるが、フレーム間の連続性を高める（円滑化する）ために、上記振幅データやピッチデータの各値を1フレーム中の例えば中心位置における各データ値とし、次のフレームの中心位置までの間（合成時の1フレーム）の各データ値を補間により求める。すなわち、合成時の1フレーム（例えば上記分析フレームの中心から次の分析フレームの中心まで）において、先端サンプル点での各データ値と終端（次の合成フレームの先端）サンプル点での各データ値とが与えられ、これらのサンプル点間の各データ値を補間により求めるようにしている。

【0113】以下、有声音合成部137における合成処理を詳細に説明する。上記V（有声音）と判別された第mバンド（第m高調波の帯域）における時間軸上の上記1合成フレーム（Lサンプル、例えば160サンプル）分の有声音を $V_m(n)$ とすると、この合成フレーム内の時間インデックス（サンプル番号）nを用いて、

る振幅データの第m高調波の値を直線補間すればよい。すなわち、上記合成フレームの先端（ $n=0$ ）での第m高調波の振幅値を A_{0m} 、該合成フレームの終端（ $n=L$ ：次の合成フレームの先端）での第m高調波の振幅値を A_{Lm} とすると、

【0115】次に、上記(15)式中の位相 $\theta_m(n)$ は、

$n=L$ における位相 ϕ_{Lm} が $\theta_m(L)$ に等しくなるような最小の $\Delta\omega$ を設定する。

【0116】以下、任意の第mバンドにおいて、それぞれ $n=0$ 、 $n=L$ のときのV/U判別結果に応じた上記振幅 $A_m(n)$ 、位相 $\theta_m(n)$ の求め方を説明する。第mバンドが、 $n=0$ 、 $n=L$ のいずれもV（有声音）と

される場合に、振幅 $A_m(n)$ は、上述した(16)式により、伝送された振幅値 A_{0m} 、 A_{Lm} を直線補間して振幅 $A_m(n)$ を算出すればよい。位相 $\theta_m(n)$ は、 $n=0$ で $\theta_m(0)=\phi_{0m}$ から $n=L$ で $\theta_m(L)=\phi_{Lm}$ となるように $\Delta\omega$ を設定する。

【0117】次に、 $n=0$ のときV(有声音)で、 $n=L$ のときUV(無声音)とされる場合に、振幅 $A_m(n)$ は、 $A_m(0)$ の伝送振幅値 A_{0m} から $A_m(L)$ で0となるように直線補間する。 $n=L$ での伝送振幅値 A_{Lm} は無声

$$\theta_m(0)=\phi_{Lm}-m(\omega_{01}+\omega_{L1})L/2 \quad \dots (18)$$

とし、かつ $\Delta\omega=0$ とする。

【0119】上記 $n=0$ 、 $n=L$ のいずれもV(有声音)とされる場合に、 $\theta_m(L)$ が ϕ_{Lm} となるように $\Delta\omega$

$$\begin{aligned} \theta_m(L) &= m\omega_{01}L + L^2 m(\omega_{L1}-\omega_{01})/2L + \phi_{0m} + \Delta\omega L \\ &= m(\omega_{01}+\omega_{L1})L/2 + \phi_{0m} + \Delta\omega L \\ &= \phi_{Lm} \end{aligned}$$

となり、これを整理すると、 $\Delta\omega$ は、

$$\Delta\omega = (\text{mod}2\pi((\phi_{Lm}-\phi_{0m}) - mL(\omega_{01}+\omega_{L1})/2)/L) \quad \dots (19)$$

となる。この(19)式で $\text{mod}2\pi(x)$ とは、 x の主値を $-\pi \sim +\pi$ の間の値で返す関数である。例えば、 $x=1.3\pi$ のとき $\text{mod}2\pi(x)=-0.7\pi$ 、 $x=2.3\pi$ のとき $\text{mod}2\pi(x)=0.3\pi$ 、 $x=-1.3\pi$ のとき $\text{mod}2\pi(x)=0.7\pi$ 、等である。

【0120】ここで、図17のAは、音声信号のスペクトルの一例を示しており、バンド番号(ハーモニクスナンバー) m が8、9、10の各バンドがUV(無声音)とされ、他のバンドはV(有声音)とされている。このV(有声音)のバンドの時間軸信号が上記有声音合成部137により合成され、UV(無声音)のバンドの時間軸信号が無声音合成部138で合成されるわけである。

【0121】以下、無声音合成部138における無声音合成処理を説明する。ホワイトノイズ発生部142からの時間軸上のホワイトノイズ信号波形を、所定の長さ(例えば256サンプル)で適当な窓関数(例えばハミング窓)により窓かけをし、STFT処理部143によりSTFT(ショートタイムフーリエ変換)処理を施すことにより、図17のBに示すようなホワイトノイズの周波数軸上のパワースペクトルを得る。このSTFT処理部143からのパワースペクトルをバンド振幅処理部144に送り、図17のCに示すように、上記UV(無声音)とされたバンド(例えば $m=8, 9, 10$)について上記振幅 $|A_m|_{UV}$ を乗算し、他のV(有声音)とされたバンドの振幅を0にする。このバンド振幅処理部144には上記振幅データ、ピッチデータ、V/UV判別データが供給されている。バンド振幅処理部144からの出力は、ISTFT処理部145に送られ、位相は元のホワイトノイズの位相を用いて逆STFT処理を施すことにより時間軸上の信号に変換する。ISTFT処理部145からの出力は、オーバーラップ加算部146

音の振幅値であり、後述する無声音合成の際に用いられる。位相 $\theta_m(n)$ は、 $\theta_m(0)=\phi_{0m}$ とし、かつ $\Delta\omega=0$ とする。

【0118】さらに、 $n=0$ のときUV(無声音)で、 $n=L$ のときV(有声音)とされる場合には、振幅 $A_m(n)$ は、 $n=0$ での振幅 $A_m(0)$ を0とし、 $n=L$ で伝送された振幅値 A_{Lm} となるように直線補間する。位相 $\theta_m(n)$ については、 $n=0$ での位相 $\theta_m(0)$ として、フレーム終端での位相値 ϕ_{Lm} を用いて、

を設定する手法について説明する。上記(17)式で、 $n=L$ と置くことにより、

に送られ、時間軸上で適当な(元の連続的なノイズ波形を復元できるように)重み付けをしながらオーバーラップ及び加算を繰り返し、連続的な時間軸波形を合成する。オーバーラップ加算部146からの出力信号が上記加算部140に送られる。

【0122】このように、各合成部137、138において合成されて時間軸上に戻された有声音部及び無声音部の各信号は、加算部140により適当な固定の混合比で加算して、出力端子141より再生された音声信号を取り出す。

【0123】ここで、上述したデコーダ側のビタビ復号&CRC検出は、以下のような原理である。図18は、ビタビ復号&CRC検出の原理を説明するための機能ブロック図である。例えば、図18に示すような原理である。まず、伝送されてきた224ビットを2スロットデインターリーブ器151が受信し、デインターリーブする。この2スロットデインターリーブ器151の出力をクラス2とエンコードされているクラス1ビットに分け、後者を畳込み復号化器152に入力し、復号して、80ビットのクラス1復号結果を受信7ビットを得る。次に、80ビットのクラス1復号結果からエンコーダで計算したのと同じパラメータビットに相当するものから再びCRCをCRC計算部153により計算し、受信CRCと比較し、その結果を音声復号器154に出力する。

【0124】なお、上記図10の音声分析側(エンコード側)の構成や図16の音声合成側(デコード側)の構成については、各部をハードウェア的に記載しているが、いわゆるDSP(デジタル信号プロセッサ)等を用いてソフトウェアプログラムにより実現することも可能である。

【0125】なお、本発明は上記実施例のみに限定されるものではなく、例えば、音声信号のみならず、音響信号を入力信号として用いることもできる。

【0126】

【発明の効果】以上の説明から明らかなように、本発明に係るデータ数変換方法によれば、ブロック内で可変個数のデータを帯域制限型オーバーサンプリングのためのFIRフィルタで、データのサンプル点に対してそれぞれ異なる複数の位相と対応した複数の係数セットの内の一定個数のデータの各位置に対応する係数セットを用いることにより、出力として必要な一定個数のデータのみを求めるので、必要な点のみを計算する間引かれた演算が可能となり、積和の演算回数を大幅に減らせる。

【0127】また、他の発明に係るデータ数変換方法によれば、ブロック内で可変個数のデータを帯域制限型オーバーサンプリングのためのFIRフィルタで、データのサンプル点に対してそれぞれ異なる複数の位相と対応した複数の係数セットの内の一定個数のデータの各位置に対応する係数セットを用いることにより、中間的な出力データを求め、この中間的な出力データを補間して最終的に必要とされる一定個数のデータを求めるので、必要な点のみを計算する間引かれた演算が可能となり、積和の演算回数を大幅に減らせる。

【0128】また、他の発明に係るデータ数変換方法によれば、ブロック内で一定個数とされたデータを帯域制限型オーバーサンプリングのためのFIRフィルタで、データのサンプル点に対してそれぞれ異なる複数の位相と対応した複数の係数セットの内のブロック内のデータ内容に応じた個数のデータの各位置に対応する係数セットを用いることにより、出力として必要なブロック内のデータ内容に応じた個数のデータのみを求めるので、必要な点のみを計算する間引かれた演算が可能となり、積和の演算回数を大幅に減らせる。

【0129】また、他の発明に係るデータ数変換方法によれば、ブロック内で一定個数とされたデータを帯域制限型オーバーサンプリングのためのFIRフィルタで、データのサンプル点に対してそれぞれ異なる複数の位相と対応した複数の係数セットの内のブロック内のデータ内容に応じた個数のデータの各位置に対応する係数セットを用いることにより、中間的な出力データを求めるので必要な点のみを計算する間引かれた演算が可能となり、積和の演算回数を大幅に減らせる。

【図面の簡単な説明】

【図1】本発明に係るデータ数変換方法の第1の実施例を説明するための概略構成を示すブロック図である。

【図2】データ数変化の一例を説明するための波形図である。

【図3】スペクトルエンベロープの拡張を説明するための波形図である。

【図4】FIRフィルタのフィルタ係数を説明するため

の図である。

【図5】図4に示されたフィルタ係数を用い実際に出力点を求める例を説明するための図である。

【図6】直線補間で使う値の求め方及び直線補間を説明するための図である。

【図7】直線補間で使う値の求め方を説明するためのフローチャートである。

【図8】直線補間を説明するためのフローチャートである。

【図9】第2の実施例を説明するための図である。

【図10】本発明に係るデータ数変換方法が適用される装置の具体例としての音声信号の合成分析符号化装置の分析側（エンコード側）の概略構成を示す機能ブロック図である。

【図11】窓かけ処理を説明するための図である。

【図12】窓かけ処理と窓関数との関係を説明するための図である。

【図13】直交変換（FFT）処理対象としての時間軸データを示す図である。

【図14】周波数軸上のスペクトルデータ、スペクトル包絡線（エンベロープ）及び励起信号のパワースペクトルを示す図である。

【図15】CRC&畳込み符号を説明するための図である。

【図16】本発明に係るデータ数変換方法が適用される装置の具体例としての音声信号の合成分析符号化装置の合成側（デコード側）の概略構成を示す機能ブロック図である。

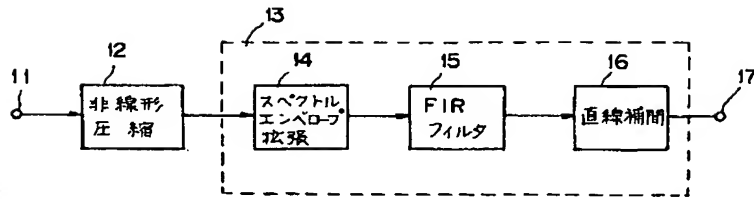
【図17】音声信号を合成する際の無声音合成を説明するための図である。

【図18】CRC&畳込み復号を説明するための図である。

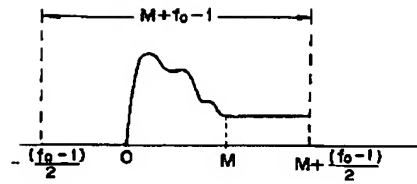
【符号の説明】

- 12・・・非線形圧縮部
- 13・・・データ個数変換本体部
- 14・・・スペクトルエンベロープ拡張部
- 15・・・帯域制限型FIRフィルタ
- 16・・・直線補間部
- 103・・・ピッチ抽出部
- 104・・・窓かけ処理部
- 105・・・直交変換（FFT）部
- 106・・・高精度（ファイン）ピッチサーチ部
- 107・・・有声音／無声音（V／UV）判別部
- 108・・・振幅再評価部
- 109・・・データ数変換（データレートコンバート）部
- 110・・・ベクトル量子化部
- 111・・・CRC&畳込み符号化部
- 112・・・フレームインターリーブ部

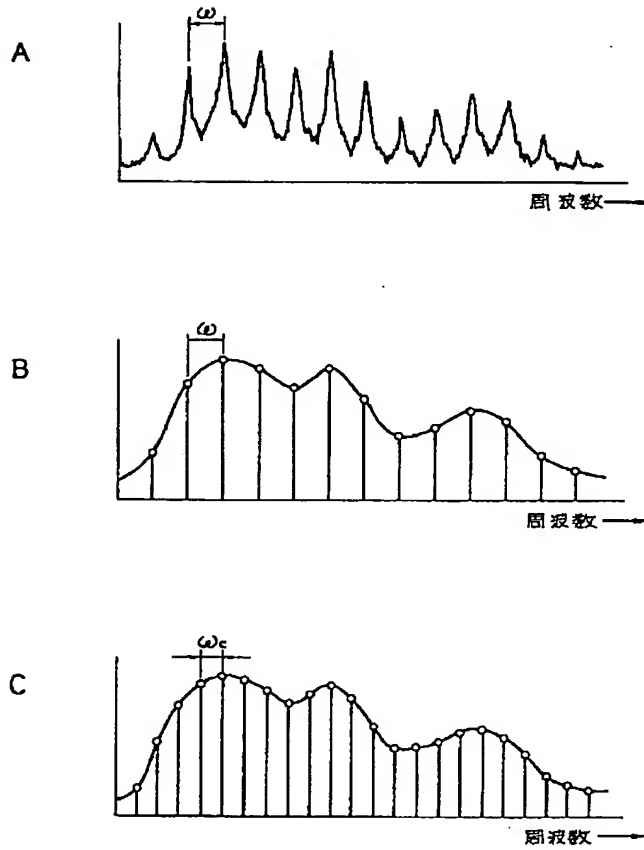
【図1】



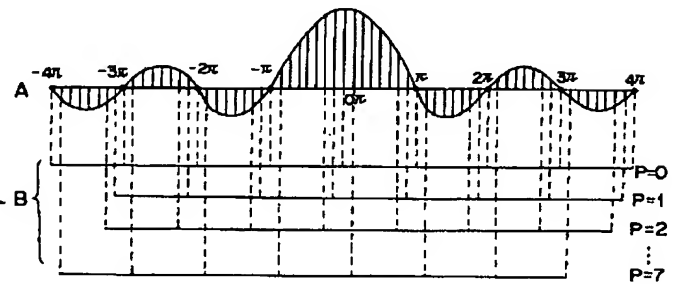
【図3】



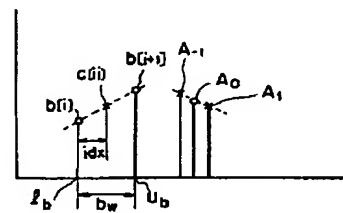
【図2】



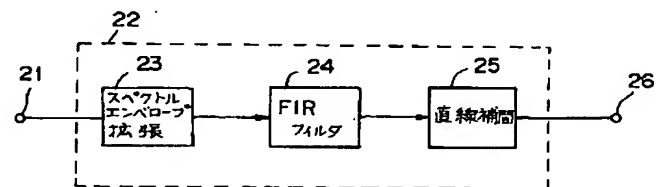
【図4】



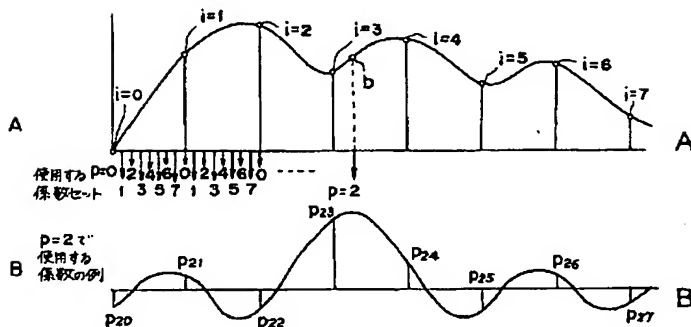
【図6】



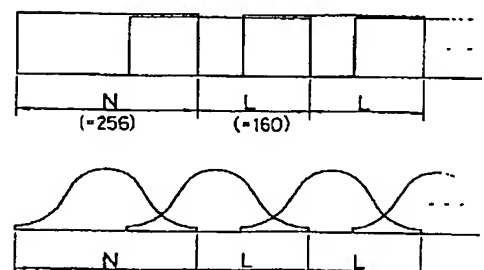
【図9】



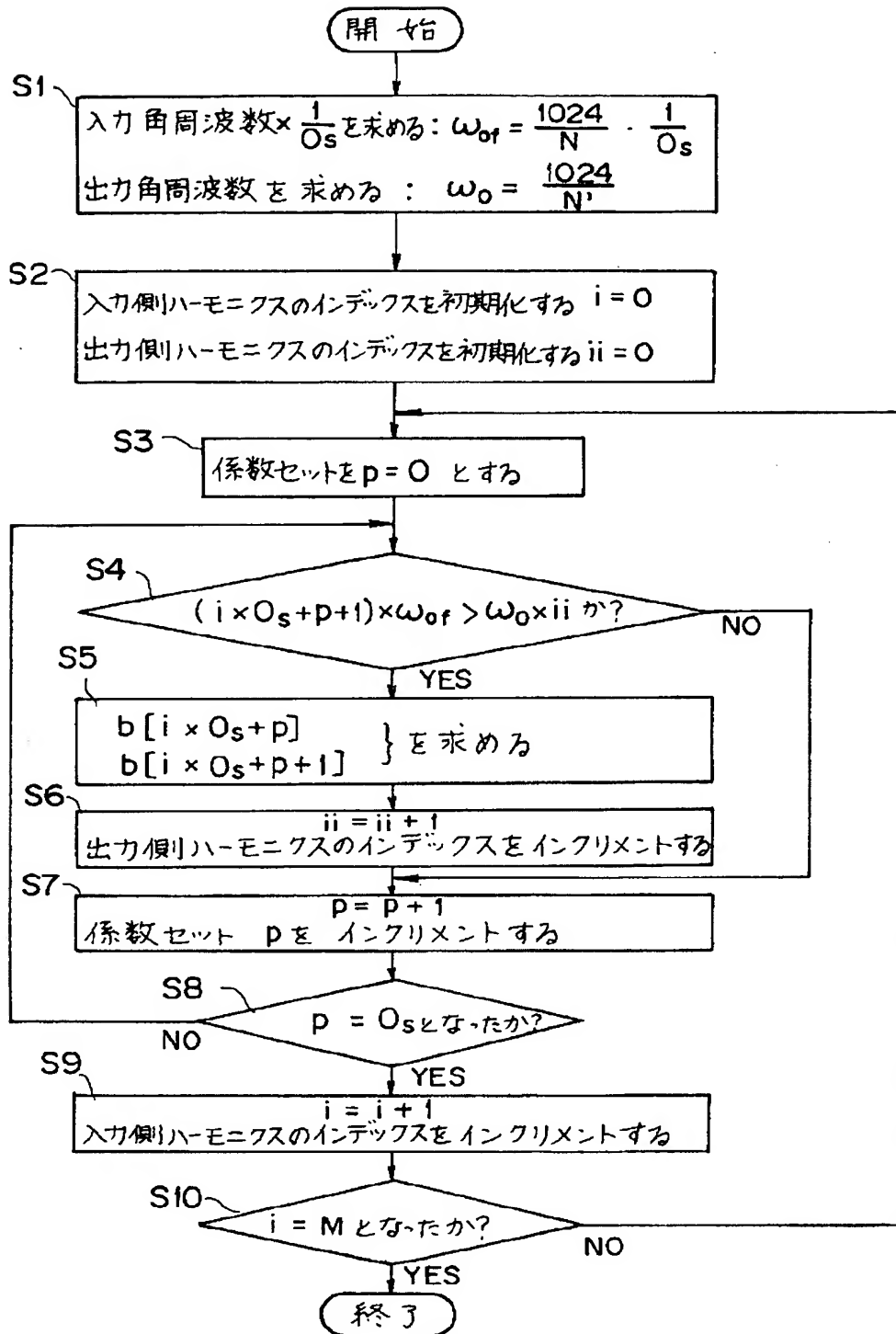
【図5】



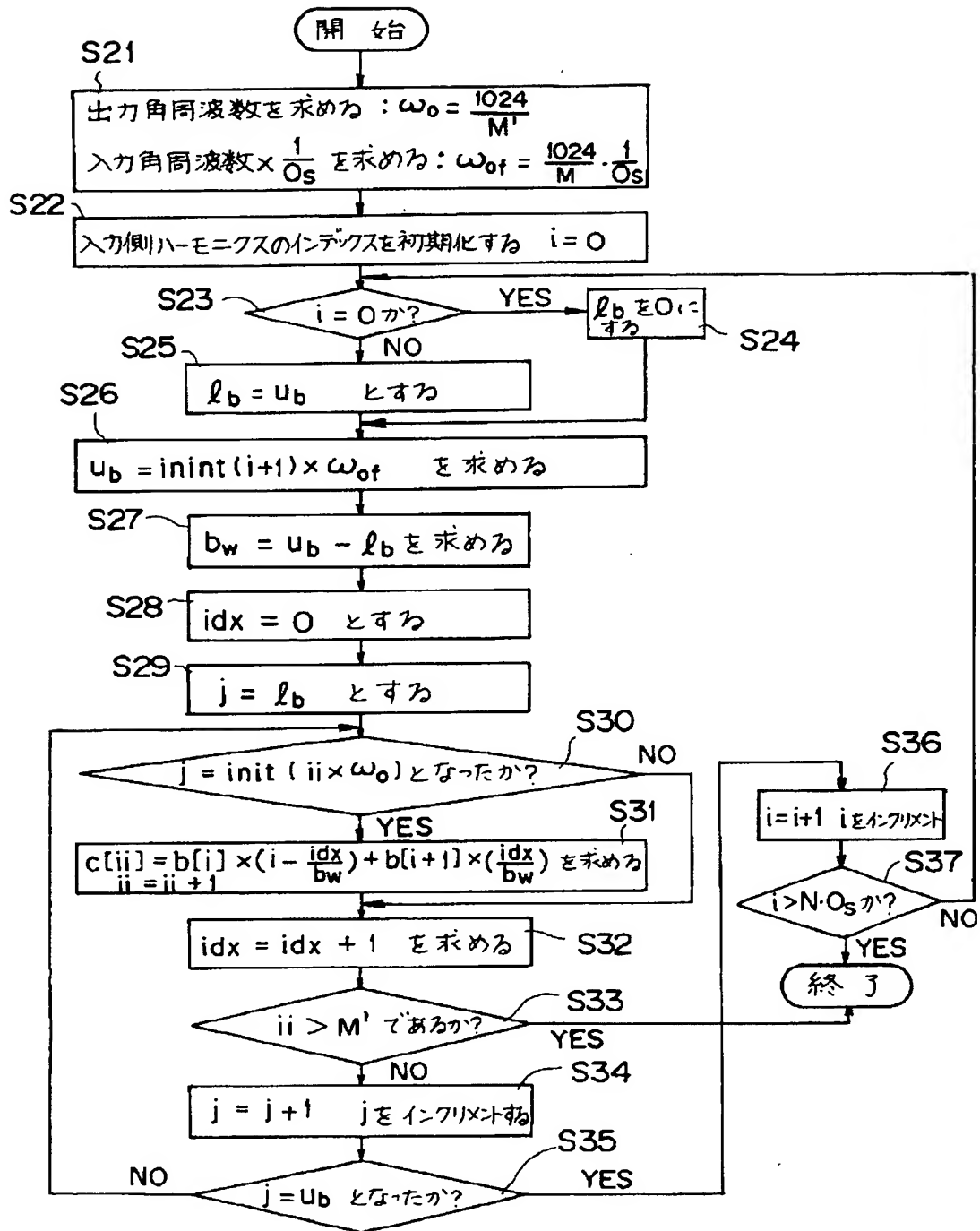
【図11】



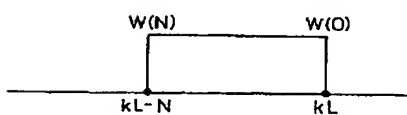
【図7】



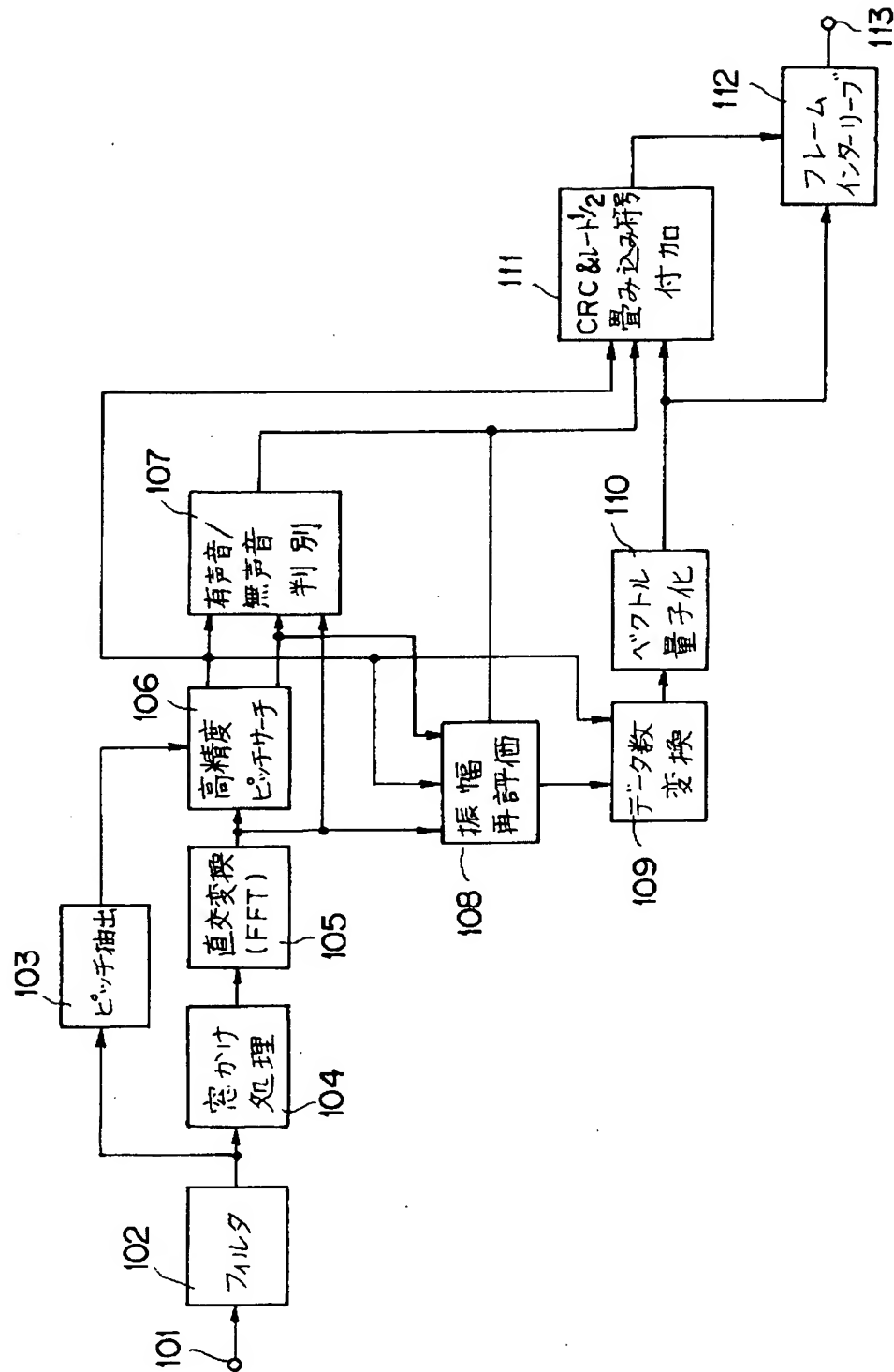
【図8】



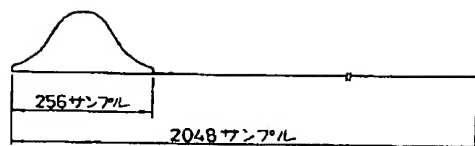
【図12】



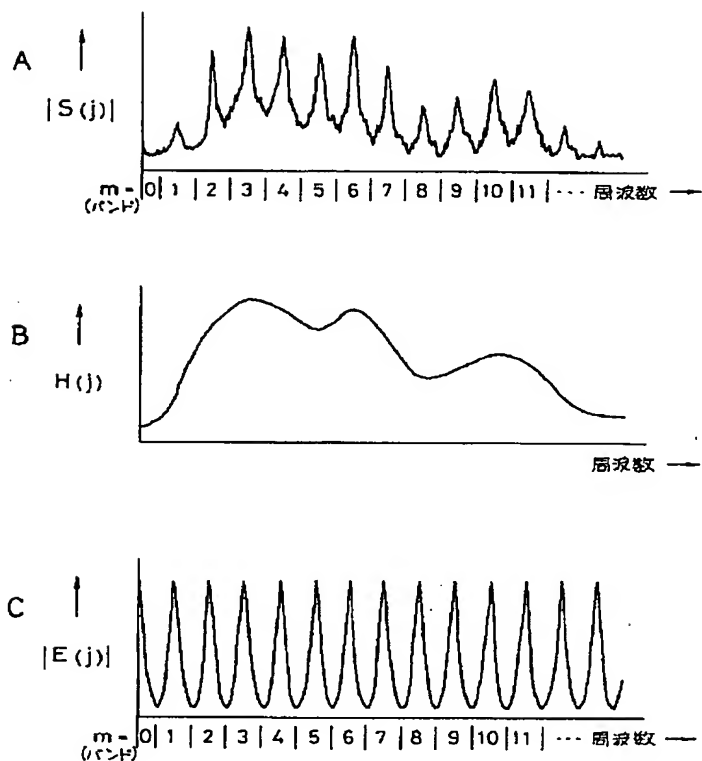
【図10】



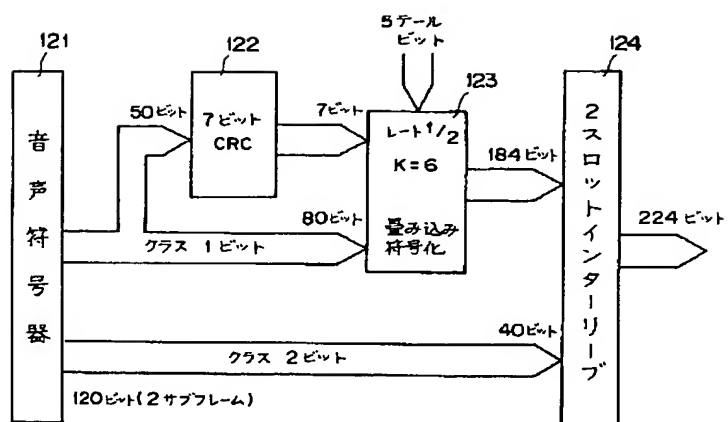
【図 13】



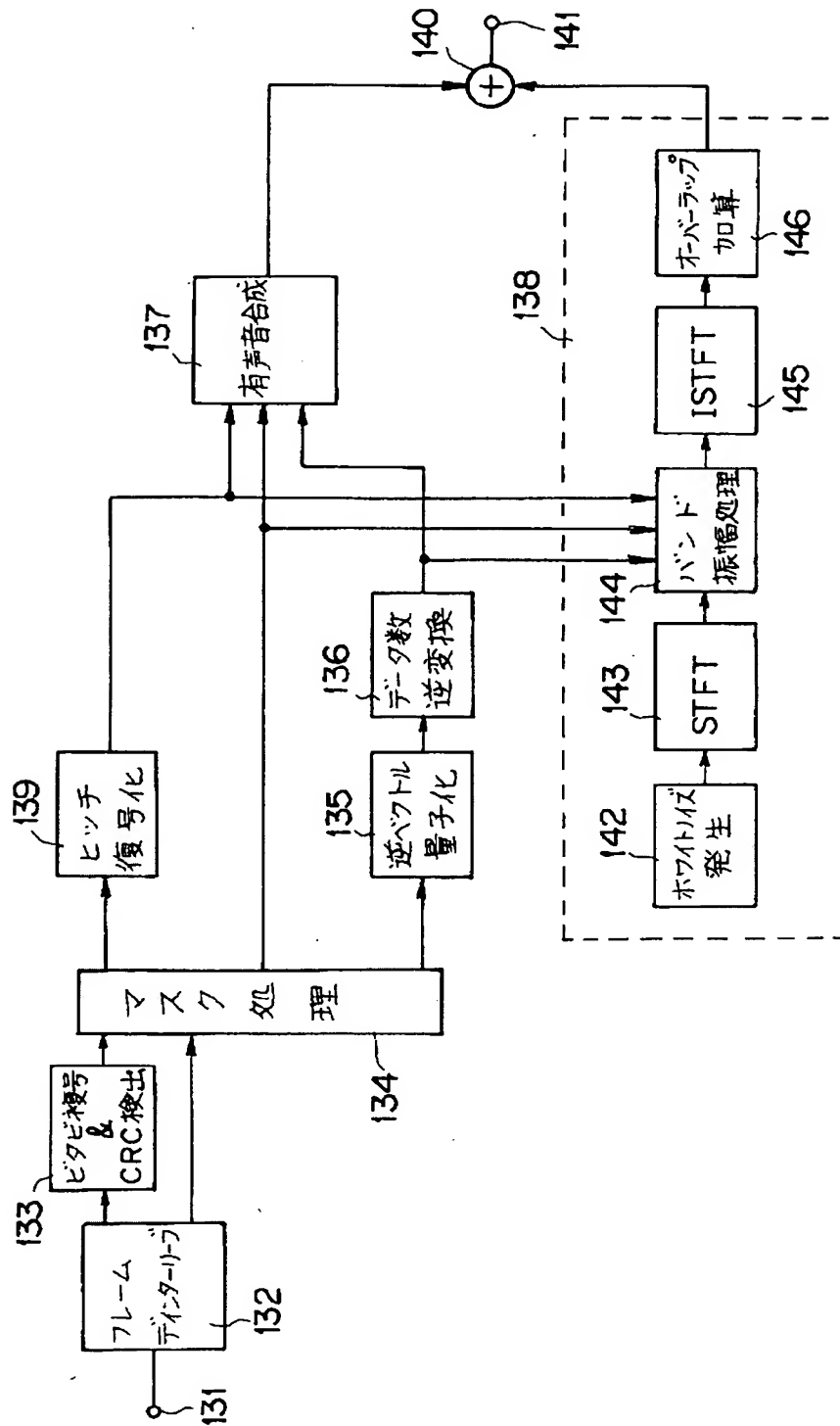
【图 14】



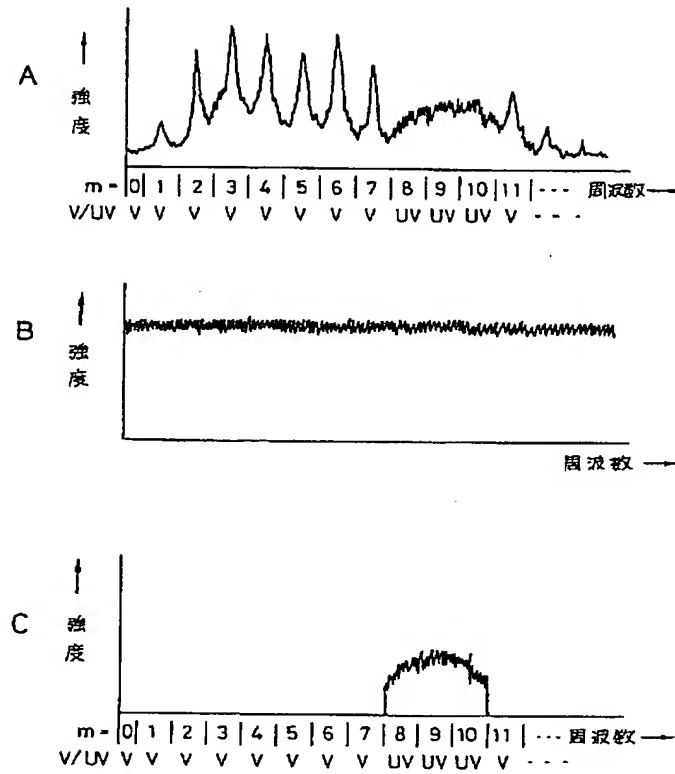
【図 15】



【図16】



【図17】



【図18】

